

Andrew Schechtman-Rook

rook166@gmail.com
(917)-836-4267

<https://github.com/AndrewRook>

TECHNICAL SKILLS

Programming: Python (numpy, scipy, pandas, sklearn, xgboost, matplotlib, pytorch), shell scripting, Javascript, C++

Databases, Orchestration, Web Design: Flask, MySQL/PostgreSQL, Prefect, Snowflake

Cloud Computing/Devops: AWS, Docker, CircleCI, Jenkins, GitHub Actions

Techniques: Model interpretability, parallel and distributed computing, Monte Carlo, bootstrapping, nonlinear optimization

Data Analysis: machine learning, visualization, experimental design

EDUCATION

PhD, Astronomy, University of Wisconsin-Madison

December 2013

MS, Astronomy, University of Wisconsin-Madison

June 2009

BS, Astronomy, Case Western Reserve University

May 2007

EXPERIENCE

Director, Data Science

2023-Present

Capital One

- Designed and built a package to streamline core business metrics SQL calculations, used by over 50 analysts for critical reporting needs.
- Owned the design, implementation, and upkeep of the DS Technical Interview, a key interview given to every data science candidate in the US.

Senior Manager, Data Science

2018-2023

Capital One

- Guided ongoing development of the core credit card valuations model scoring platform, delivering regular releases of new and updated models while improving the robustness and maintainability of platform infrastructure.
- Led technical development of model monitoring tools, mentoring three junior data scientists to deliver a maintainable package on time and to spec.
- Deployed the first cloud-based credit card underwriting model in the company via a dockerized Python API, with an estimated incremental value of \$35MM per year.

Manager, Data Science

2016-2018

Capital One

- Led development of a prototype language-agnostic automated machine learning model deployment framework for cloud-based applications, influencing the development direction for the company-wide credit card application processing platform.
- Created the longest-lived, most successful internal data science tool in the company, used in production models by dozens of data scientists across multiple lines of business.

Principal Data Scientist

2014-2016

Capital One Labs

- Implemented a novel approach to deliver internal technical trainings, providing over 5000 hours of classes with no instructors.
- Programmed and deployed an interactive course completion dashboard using Flask and dc.js to provide progress reports to individual students as well as company leadership.

Postdoctoral Research Associate

2014

University of Wisconsin-Madison

- Devised metrics to improve correspondence between numerical models and astronomical data. Implemented in highly optimized Python, was able to refine agreement by up to 20% with minimal increase in computation time.
- Built a fast Voronoi Tessellation algorithm to adaptively bin images, preserving spatial resolution while maximizing signal in images with over one million pixels.
- Trained and mentored undergraduate and graduate students in programming, data analysis and statistical methods.

Research Assistant

2007-2013

University of Wisconsin-Madison

- Developed non-linear Levenberg-Marquardt χ^2 fitting algorithms using a combination of Python and C++ to constrain models of spiral galaxies to data.
- Employed on-campus distributed computing resources to perform large-scale modeling in parallel, using over 20 years of computer time in 1 month.
- Assembled a hybrid C++/Python processing pipeline to clean, register, and mosaic thousands of high-resolution images with minimal user intervention, resulting in a factor of 10+ increase in analysis precision.
- Created a genetic algorithm in C++ to efficiently fit galaxy models with unusually large numbers of free parameters to high-resolution images.